

# 웨이브릿 변환과 인간 시각 특성을 고려한 비디오 압축

김진태<sup>\*</sup>, 김동욱<sup>\*\*</sup>, 고덕영<sup>\*\*\*</sup>

<sup>\*</sup>한서대학교 컴퓨터학과, <sup>\*\*</sup>전주대학교 전기전자정보통신공학부

<sup>\*\*\*</sup>전주공업대학 전자공학과

## Video Compression Using Wavelet Transform and Human Visual Characteristics

Jin-Tae Kim<sup>\*</sup>, Dong-Wook Kim<sup>\*\*</sup>, Duck-Young Ko<sup>\*\*\*</sup>

<sup>\*</sup>Hanseo University · <sup>\*\*</sup>Jeonju University, <sup>\*\*\*</sup>Chonju Technical College

E-mail : jtkim@gaya.hanseo.ac.kr

### 요 약

본 논문에서는 움직임 보상과 웨이브릿 변환 부호화 기법을 이용하여 비디오 신호의 데이터를 압축하고자 한다. 비디오 신호에 대한 인간 시각 특성을 분석하여 큰 이득을 주지 못하는 과정을 생략하고 인간의 눈에 민감하지 않은 정보를 생략함으로써 화질의 저하 없이 비디오 신호를 높은 비율로 압축한다. 제안한 방법은 움직임이 발생하는 정도와 크기에 매우 의존적으로, 움직임이 많은 비디오 신호에 더욱 적합하나 움직임이 작은 경우에도 저주파 대역들의 움직임 보상에 의한 이득이 크기 때문에 모든 영상에 대해서 적용 가능한 방법이다. 'Football'과 'Popple'을 대상으로 한 컴퓨터 시뮬레이션 결과, 프레임당 30:1 정도의 압축하에서 높은 재생화질을 얻을 수 있었다.

### I. 서 론

디지털 전송기술, 영상처리 기술 등의 발전과 더불어 영상통신 분야는 차세대 통신의 중요한 분야가 되었다. MPEG-1,2의 표준화가 이루어졌고, HDTV는 21세기에 상용화가 될 예정이다. 디지털 비디오 신호는 대역폭이 매우 넓어 저장 및 전송을 하기 위해서는 고압축을 하지 않으면 안 된다. 예를 들면, 1280×720 크기의 비디오 신호 60프레임을 1초당 전송하는데 할당된 대역폭은 20~22M비트 정도이므로 30:1 이상의 압축을 요하는 것이다.[1]

웨이브릿(wavelet) 이론은 다양한 스케일(scale)과 해상도(resolution)에서 신호를 고찰, 분석하고자 하는 생각으로 수학, 공학분야에서 발전되어 왔다.[2] 이러한 시도는 1980년대 중반 Morlet, Grossmann, Meyer 등에 의해 개괄적인 수학 체계가 이루어졌으며, 이후 Doubechies와 Mallat에 의해 이산 신호처리로 확대, 발전되었다. 웨이브릿 변환을 영상처리에 적용하면, 영상을 특정 주파수 대역으로 나누어 처리할 수 있고, 영상신호를 다해상도 계층 구조화하는데 많은 이점이 있다.[3] 따라서 근래에 영상처리 분야에 웨이브릿 변환이 많이 사용되고 있다.

본 논문에서는 TV 신호와 비디오와 같은 영상 신호(sequence images)를 웨이브릿 변환을 사용하여 고화질을 유지한 고압축을 이루고자 한다. 웨이브릿 변환을 이용하여 영상신호의 대역을 압

축하고자할 때, 대역 세분에 의한 공간영역상의 압축으로는 충분한 압축을 이룰 수 없으므로 시간축상의 중복성을 이용하는 방법이 도입되고 있다. 웨이브릿 변환에 움직임 보상을 함으로써 시간축상의 중복성을 이용한 데이터 압축을 이룰 수 있다.[4] 본 논문에서는 영상을 16개의 등대역으로 분할한다. 4개의 저주파 대역들은 에너지가 높고 재생화질에 지대한 영향을 미치는 대역들이므로 움직임 보상을 하고 DCT하여 에너지 압축을 행한다. 나머지 12개 대역에 대해서는 저주파 대역의 움직임 벡터 크기에 근거하여 인간 시각 특성을 고려한 처리를 한다. 고주파 대역이 갖고 있는 성분들은 대체로 현재 프레임의 에지 성분들이다. 이들 에지 성분들은 해당 부위의 움직임의 크기에 따라서 중요도가 결정되고, 이렇게 함으로써 움직임이 작거나 정지한 영역에 대해서는 그 선명도를 살리면서 움직임이 큰 영역에 대해서는 시각적 지장을 주지 않는 범위내에서 데이터를 크게 줄일 수 있다. 또한 제안한 방법은 움직임이 발생하는 정도와 크기에 매우 의존적이거나, 움직임이 작은 영상의 경우에는 고주파 대역에서의 이득은 낮으나 4개의 저주파 대역들의 움직임 보상에 의한 이득이 크기 때문에 모든 영상에 대해서 적용 가능한 방법이다.

### II. 인간 시각 특성

영상신호의 한 프레임은 시간축에서 샘플링된

일정한 해상도를 가진 영상 데이터로서, 영상의 밝기와 명확도에 따라서 인간의 눈이 느끼는 감도가 다르다.

비디오와 같이 프레임율(frame rate)이 높아 어떤 밝기에서도 인간의 시각이 영상의 시간축에서 불연속을 느끼지 못할 경우에는 영상 내의 영역의 상세한(detail) 정도에 따라서 인간 시각의 감도가 다르다. 그림 1에서 볼 수 있는 바와 같이, 인간의 눈은 시간축으로 저역통과 필터(low-pass filter)의 특성을 갖고, 그 필터링 효과는 낮은 공간 주파수(평탄한 영역)보다는 높은 공간 주파수(섬세한 영역)에 대해서 더 크게 나타난다. 즉, 높은 공간 주파수 영역은 낮은 프레임율로 전송되어도 인간의 눈은 거의 감지하지 못한다는 것을 의미하고, 이를 이용하여 효율적인 전송을 할 수 있다.[5] 영상신호에 대한 인간 시각 특성을 정량적으로 표현하기는 어려운 일이나, 인간 시각 특성의 영상신호 대역압축에의 적용은 인간의 눈이 감지할 수 없는 정보를 제거할 수 있게 한다.

그림 2는 영상신호에 대한 인간 시각 특성을 공간 주파수 영역에서 표시한 것이다. 공간 주파수 영역상의 인간 시각 특성은 대역통과 필터(band-pass filter)의 특성을 갖고 있으며, 영상내 각 부분의 움직임 정도에 따라서 피크(peak)의 크기와 통과 대역을 달리한다. 움직임의 속도가 클수록 피크의 위치는 영(zero) 주파수쪽으로 이동하고 상대적 감도도 크게 떨어지는 것을 볼 수 있다.[6] 다시 말하면, 영상 내의 움직임이 큰 부분은 어느 정도 거친 처리를 하여도 인간 시각에는 별로 영향을 주지 않는다. 이러한 인간 시각 특성의 이용은 주관적 고품질을 유지한 채로 영상신호를 고압축할 수 있는 가능성을 제시한다.

### III. 시각 특성을 이용한 영상 압축

시각 특성을 영상신호의 압축에 이용한 기존의 연구들이 정지영상을 대상으로 각 주파수 영역에 대한 인간 시각의 상대적 감도(relative sensitivity)에 관심을 둔 반면에 본 논문에서는 일정한 프레임율을 갖는 비디오 신호에 대해서 인간 시각 특성을 적용하였다.

각 프레임을 그림 3과 같이 16개의 등대역으로 분할하였다. B-11,12,21,22 대역들은 재생영상의 화질에 매우 큰 영향을 미치는 대역이고, 나머지 12개의 대역들은 높은 공간 주파수 성분들로 구성된 대역 영상들로서 비교적 작은 영향을 미친다. 앞으로 이들 대역들을 각각 저주파수 대역과 고주파수 대역으로 구분하여 칭하겠다. 고주파수 대역들에 대해서 인간 시각의 시간축상의 저역통과 필터 특성과 공간주파수 영역에서의 움직임 속도에 따른 인간 시각의 대역통과 필터 특성을 적용할 수 있다. 즉, 영상은 움직이는 영역과 배경 영역, 배경 영역은 평탄한 영역과 세밀한 영역

으로 구분되어 처리될 필요가 있다. 또한 시각적 중요도 측면에서 저주파수 대역은 고주파수 대역보다 충실히 처리해야 한다. 이렇게 함으로써 영상의 각 영역이 일정한 화질을 유지하게 되고, 전체적으로는 주관적 고품질을 유지할 수 있게 된다.

영상신호는 프레임내 처리를 하는 프레임과 프레임간 처리를 하는 프레임으로 나눌 수 있다. 각각을 *intraframe*과 *interframe*으로 부르겠다. 장면이 바뀔 때의 첫 번째 프레임과 그후 몇 프레임중 한 프레임은 *intraframe* 처리된다. *Intraframe*의 복원화질은 연이어지는 다음 프레임들(*interframes*)의 복원화질을 기본적으로 좌우한다.

*Intraframe* 처리의 경우(그림 4(a)), B-11,12,21,22의 저주파수 대역이 다른 대역에 비해 월등히 높은 에너지를 갖고 있고, 복원화질을 크게 좌우하므로 신호성분의 손실을 최소로 하는 압축기법을 채택하여야 한다. 따라서 이 대역들에 DCT를 적용하였다. B-11,12,21,22를 제외한 나머지 고주파수 대역은 인간 시각의 저역통과 필터 특성을 적용하여 거칠게 처리되어도 주관적 화질에 큰 영향을 주지 않는다. B-11,12,21,22 대역의 DCT 계수들은 좁은 양자화 계단폭을 갖는 균일 양자화기로 양자화된다. 나머지 대역들은 에너지 압축을 하지 않고 바로 균일 양자화되며, 양자화 계단폭의 조절에 의해 일정한 전송율을 유지하게 된다. 최종 양자화된 대역영상은 VLC(variable length coding)된다.

*Interframe*에서는(그림 4(b)) 시간축으로 나타나는 중복성을 이용하기 위해서 저주파수 대역에 움직임 보상을 한다. 즉, B-11,12,21,22 대역신호가 움직임 보상에 의한 예측 오차인 반면 나머지 대역신호들은 현재 프레임에 대한 고주파 성분들이고 시간축상의 변화에 대한 정보를 포함하고 있지 않다. 영상신호에 대한 인간 시각 특성의 적용 측면에서, 이들 고주파수 대역의 데이터 압축은 저주파수 대역에서 구한 움직임 벡터(움직임의 크기)에만 의존한다. 각 위치의 움직임 벡터의 크기에 따라서 움직임 영역과 배경 영역으로 구분된다. 각 영역별로 유효한 대역을 달리하여 유효하지 않은 대역의 해당위치 신호성분을 제거한다. 저주파수 대역신호는 DCT되어 좁은 양자화 계단폭을 갖는 균일 양자화기로 양자화되고, 고주파수 대역신호는 바로 균일 양자화되며, 양자화 계단폭의 조절에 의한 일정한 전송율을 유지하게 된다. 최종 양자화된 대역영상은 VLC된다.

### IV. 컴퓨터 시뮬레이션 및 검토

대상영상으로는 MPEG 지정 표준영상인 'Football'과 'Popple'을 사용하였다. 'Football'은 움직임이 상당히 크고 배경은 전체적으로 고르며 고주파 성분을 많이 포함한 영상이며, 'Popple'은

배경 부분이 매우 평탄하며 극히 저주파 성분으로 이루어져 있고 고정된 위치에서의 회전 운동만이 있는 영상이다. 대상영상 신호들은 4:2:2의 디지털 신호로 휘도신호(Y)와 색차신호(Cb,Cr)로 구성되어 있고 비월주사된 형태이다. 본 실험에서는 휘도신호만을 대상으로 홀수번째 필드를 세로축으로 보간하여 순차주사된 형태로 바꾸어 그것을 한 프레임으로 하였다. 영상의 크기는 720×480이고, 영상신호의 프레임율은 30frames/sec이고, 한 프레임의 intraframe처리 후에 9 프레임의 interframe처리가 뒤따른다. 표 1은 본 실험에서 사용한 웨이브릿 필터 계수이고, Kernel은 0.6이다.

한 프레임의 대역 신호는 저주파수 대역신호와 고주파수 대역신호로 나눌 수가 있다. 영상의 16-대역분할을 기본으로 하여 시뮬레이션 모델 SM1을 설정하였다. SM1에서 저주파수 대역의 움직임 보상은 B-11의 움직임 벡터를 B-12,21,22에 그대로 적용하는 방법을 채택했다. 이때, B-11 대역신호는 원영상에 비해 가로·세로 크기의 1/4 해상도를 가지므로 움직임 추정 블록의 크기를 4×4, 움직임 추정부 변위를 ±3으로하여 BMA의 full search 방법을 사용하였다. 또한 블록 경계의 불연속으로 인한 부수적 고주파 성분의 발생을 최소화하기 위해서 DCT 블록의 크기도 4×4로 같게 했다. 고주파수 대역에서도 마찬가지로 4×4 블록단위로 각 대역 해당위치의 중요도가 B-11의 움직임 벡터에 의해 결정된다. 그러나 SM1의 저주파수 대역의 경우, 4×4의 작은 블록에 대해서 DCT를 적용하였기 때문에 블록의 크기가 작은 만큼 DCT의 압축 효율이 떨어지고 모든 DCT 계수가 복원된 영상에 민감하게 영향을 미칠 수 있다.

이에 대한 대안으로 시뮬레이션 모델 SM2를 설정하였다. SM2에서는 저주파수 대역영상을 원영상의 가로·세로 크기의 1/2 해상도를 유지하게 하였다. 따라서 움직임 추정 및 DCT 블록의 크기를 8×8로 하였다. 이 대역의 움직임 벡터 크기를 1/2로하여 SM1의 고주파수 대역 처리에 사용될 수 있다.

시뮬레이션 모델 SM1은 고주파수 대역의 좀더 높은 비율로의 데이터 압축을 위한 것이다. 실제로 양자화된 후의 고주파수 대역의 많은 부분은 영(zero)의 값을 갖으나, 블록 단위의 정보전송으로 인해 불필요한 비트가 할당되고 있다. B-23,32,33,34,43,44의 대역들을 전송하지 않음으로써 고주파수 대역에서의 데이터를 압축할 수 있다. 시뮬레이션 모델들에 대해서 그림5에 나타내었다. 영상처리 방법은 시뮬레이션 모델(SM1, SM2, SM3)의 혼성 형태를 취하였다. 아래 표현의 L과 H는 각각 저주파수 대역과 고주파수 대역을 나타낸다.

(방법1)SM1,L+SM1,H (방법2)SM2,L+SM1,H

(방법3)SM1,L+SM3,H (방법4)SM2,L+SM3,H

각 시뮬레이션 모델의 성능을 비교하기 위해서 저주파수 대역과 고주파수 대역의 양자화 계단폭을 각각 2와 8로 고정시켜 놓고 각 대역을 전송하는데 필요한 비트를 구하였다.(그림 6) 그림 7은 (방법1)-(방법4)로 영상신호를 처리한 결과이다.

처리된 영상은 전반적으로 높은 PSNR과 주관적으로도 좋은 화질을 나타내었다. 저주파수 대역과 고주파수 대역의 비트 할당 결과는 대상영상의 특징에 따라서 차이가 있었다. 'Football' 영상의 경우, (방법1)을 사용하였을 때는 저주파 대역과 고주파 대역에 각각 약 3:1의 비트 할당이 되었고 전체적인 압축율은 21:1 정도였다. (방법3)을 사용하면 전체 압축율은 24:1 정도가 된다. 'Popple'의 경우, (방법1)에서는 저주파 대역과 고주파 대역의 비트 할당 비율이 거의 비슷하며 전체 압축율은 약 27:1이었다. (방법3)을 적용한 결과는 고주파 대역에 거의 이득을 주지 못했으며 움직임 블록이 매우 적은 영상이므로 일부 고주파수 대역의 제거는 PSNR의 저하를 가져왔다. 그 이유는 영상이 포함하고 있는 공간 주파수 성분이 매우 적기 때문이라 볼 수 있다.

대상영상에 저주파 대역에 (방법2)를 적용한 결과는 (방법1)의 사용결과와 비슷하였다. DCT 블록 크기의 확대로 더 큰 압축 효과를 기대했으나 저주파수 대역의 대역 세분으로 인한 압축 효율을 증가하지 못했다. 또한 예측 오차를 DCT하기 때문에 DCT 블록 크기에 따른 영향은 매우 적다고 할 수 있다. 본 시뮬레이션에서는 균일 양자화기를 사용하여 일정범위의 화질을 유지한채 양자화 계단폭을 조정하는 것에 큰 제한을 받았는데 불균일 양자화기를 사용하면 이런 문제점은 어느 정도 해결되고 압축율도 크게 높일 수 있을 것이라고 생각한다. 그림 7에서 볼 수 있는바와 같이, PSNR은 영상마다 차이가 있으나 각 영상의 특징에 따른 처리를 했기 때문에 매우 좋은 주관적인 화질을 나타냈고 전 프레임에 걸쳐 고른 화질을 유지했다.

## V. 결론

본 논문에서는 비디오 신호에의 적용을 목적으로, 영상신호를 고압축할 수 있는 기법을 연구하였다. 웨이브릿 변환은 인간 시각 특성을 적용하기가 매우 용이하다. 인간의 눈은 시간축상으로 저역통과 필터의 특성을 갖고, 공간 영역상으로는 대역통과 필터의 특성을 갖는다. 이를 이용하여, 각 대역별로 일률적인 처리를 하지 않고 움직임 정보를 이용하여 움직임 정도에 따라 적응적인 처리를 함으로써 주관적인 화질의 열화없이 고압축을 할 수 있었다. 저주파 대역과 고주파 대역으

측을 할 수 있었다. 저주파 대역과 고주파 대역으로 구분하여 시뮬레이션 모델을 설정함으로써 본 방법의 개선 가능성을 보였다. 또한 동영상에 대상으로한 처리만큼 재생된 신호의 화질 평가에 있어서도 이를 고려한 일반적인 기준이 필요함을 느꼈다.

### 참고문헌

- [1] M. Rabbani and P. W. Jones, *Digital Image Compression Techniques*, SPIE Optical Engineering Press, 1991.
- [2] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," *IEEE Signal Processing Mag.*, pp. 14-38, Oct. 1991.
- [3] S. G. Mallat, "Multiresolution Approximations and Wavelet Orthogonal Bases of  $L^2(\mathbb{R})$ ," *Tran. Amer. Math. Soc.*, vol. 315, no. 1, pp. 69-87, Sep. 1989.
- [4] 김진태, 장준필, 김동욱, 최종수, "웨이브릿 변환된 다해상도 영상을 이용한 계층적 움직임 추정", *대한전자공학회 논문지-B*, vol. 33-B, no. 8, pp. 50, 59, 1996.
- [5] A. N. Netravali and B. G. Haskell, *Digital Pictures*, Plenum Press, 1988.9
- [6] Y. Ninomiya et al, "Present Status of Muse," *Proc. 2nd Int. Workshop on Signal Processing of HDTV*, pp. 579-602, 1988.

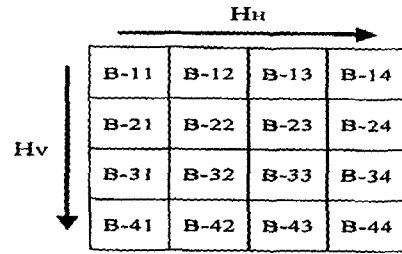
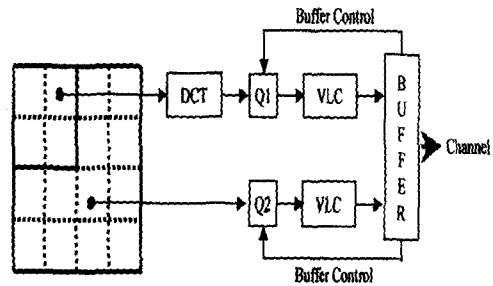
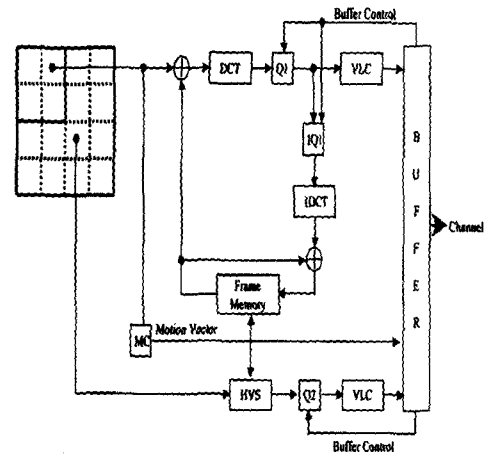


그림 3. 영상의 웨이브릿 변환



(a) Intraframe 처리



(b) Interframe 처리

그림 4. 영상처리의 블록도

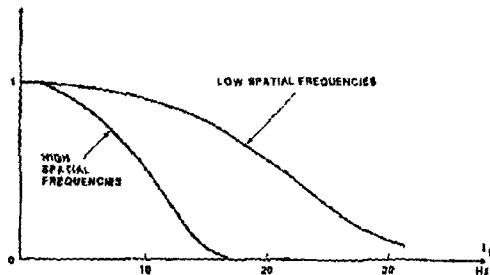


그림 1. 시간 주파수 영역에서 인간 시각 특성

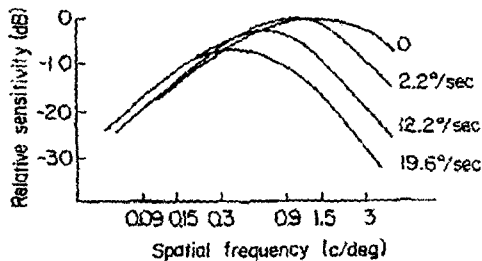
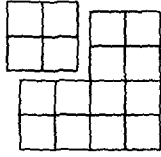
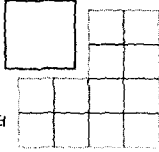


그림 2. 공간 주파수 영역에서 인간 시각 특성

[시뮬레이션 모델, SM1]  
 움직임 추정 블록 크기 : 4x4  
 움직임 추정 범위 : -3 ~ +3  
 DCT 블록 크기 : 4x4  
 움직임 정보 : B-11의 움직임 벡터



[시뮬레이션 모델, SM2]  
 움직임 추정 블록 크기 : 8x8  
 움직임 추정 범위 : -7 ~ +7  
 DCT 블록 크기 : 8x8  
 움직임 정보 : 저주파수 대역의 움직임 벡터



[시뮬레이션 모델, SM3]  
 전송되는 고주파수 대역 :  
 B-13, 14, 24, 31, 41, 42  
 전송되지 않는 고주파수 대역 :  
 B-23, 32, 33, 34, 43, 44  
 (동영상 부분)

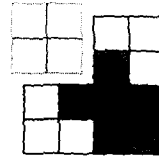
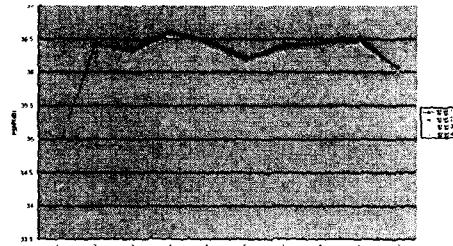
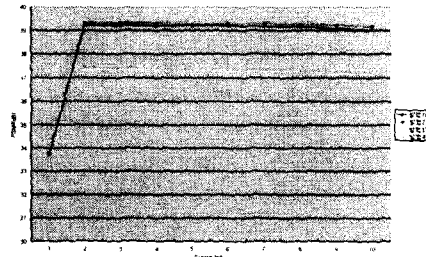


그림 5. 시뮬레이션 모델

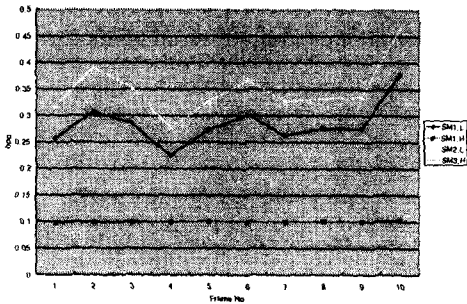


(a)'Football' 영상

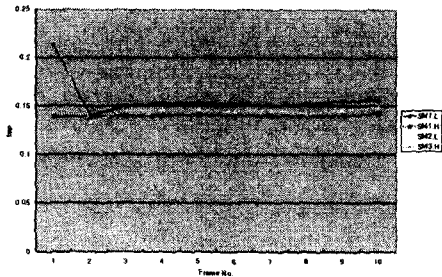


(b)'Popple' 영상

그림 7. 복원 화질



(a)'Football' 영상



(b)'Popple' 영상

그림 6. 전송율

표 1. 웨이브릿 필터 계수

분해 $h'(n)$		합성 $h(n)$	
$h'(-2)$	-0.05	$h(-3)$	-0.0107
$h'(-1)$	0.25	$h(-2)$	0.05357
$h'(0)$	0.6	$h(-1)$	0.2607
$h'(1)$	0.25	$h(0)$	-0.607
$h'(2)$	-0.05	$h(1)$	0.2607
		$h(2)$	0.05357
		$h(3)$	-0.0107